



Title	Development of behavioural automaticity by extended Pavlovian training in an insect
Author(s)	Mizunami, Makoto; Hirohata, Sho; Sato, Ai; Arai, Ryoichi; Terao, Kanta; Sato, Misato; Matsumoto, Yukihiisa
Citation	Proceedings of the royal society b-biological sciences, 286(1894), 20182132 https://doi.org/10.1098/rspb.2018.2132
Issue Date	2019-01-02
Doc URL	http://hdl.handle.net/2115/75706
Type	article (author version)
File Information	Mizunami PRSB 20182132.pdf



[Instructions for use](#)

Development of behavioral automaticity by extended Pavlovian training in an insect

Makoto Mizunami¹, Sho Hirohata², Ai Sato³, Ryoichi Arai⁴, Kanta Terao², Misato Sato²,
Yukihiisa Matsumoto^{1,5}

¹Faculty of Science, Hokkaido University, Sapporo 060-0810, Japan, ²Graduate School of Life Science, Hokkaido University, Sapporo 060-0810, Japan, ³Graduate School of Life Sciences, Tohoku University, Sendai 980-8577, Japan; ⁴Department of Biological Sciences, Hokkaido University, Sapporo, 060-0810, Japan, ⁵College of Liberal Arts and Science, Tokyo Medical and Dental University, Ichikawa, 272-0827, Japan.

Correspondence to: Prof. Makoto Mizunami, Faculty of Science, Hokkaido University, Sapporo 060-0810, Japan, mizunami@sci.hokudai.ac.jp

Key words: habit formation, classical conditioning, octopamine, cricket, reward devaluation

Running head: “habit formation” by Pavlovian training

Abstract

The effect of repetitive training on learned actions has been a major subject in behavioral neuroscience. Many studies in instrumental conditioning in mammals, including humans, suggested that learned actions early in training are goal-driven and controlled by outcome expectancy, but they become more automatic and insensitive to reduction in the value of the outcome after extended training. It was unknown, however, whether development of value-insensitive behavior also occurs by extended training of Pavlovian conditioning in any animals. Here we show that crickets *Gryllus bimaculatus* that had received minimal training to associate an odor with water (unconditioned stimulus, US) did not exhibit conditioned response (CR) to the odor when they were given water until satiation before the test, but those that had received extended training exhibited CR even when they were satiated with water. Further pharmacological experiments suggested that octopamine neurons, the invertebrate counterparts of noradrenaline neurons, mediate US value signals and control execution of CR after minimal training, but the control diminishes with the progress of training and hence the CR becomes insensitive to US devaluation. The results suggest that repetitive sensory experiences can lead to a change from a goal-driven response to a more automatic one in crickets.

1. Introduction

Many studies in mammals, including humans, suggest that learned actions early in instrumental training are goal-directed and controlled by outcome expectancy, but similar behaviors often proceed as automatic responses to antecedent stimuli with repetition of training [1-3]. Such a change, called habit formation, has been demonstrated using a post-training outcome devaluation procedure, in which the value of the outcome (food) is decreased by either taste-aversion learning or satiation: action early in training is, in large part, sensitive to devaluation but the sensitivity decreases with extended training [1,2]. The development of behavioral autonomy with extended training, however, is not found in some experimental situations, for example, when animals are subjected to complex conditioning tasks [1.4.5]. Therefore, the number of training trials is not the only factor to determine the development of behavioral autonomy. In humans, this shift can be beneficial in everyday life but can also lead to loss of control over compulsive behavior, namely drug seeking in addiction [6-8]. Lesion, electrophysiological and transgenic studies in rodents suggested that the transition from goal-oriented to habitual behavior involves transitions in the neural circuits controlling the behaviors [7,9,10], but the cellular and molecular mechanisms of habit formation remain largely unknown.

The effect of extended training has also been studied in Pavlovian conditioning in rats, in which they learn the relationship between a previously neutral stimulus (conditioned stimulus, CS) and a stimulus with biological value (unconditioned stimulus, US) and produce response (conditioned response, CR) to the CS thereafter. After minimal training, animals that received a procedure to decrease the value of the US exhibited a decrement of the CR, and hence it has been argued that the CR is controlled by CS-induced activations of representations of the US, analogous to that actions are driven by representations of goals in instrumental conditioning. In contrast to instrumental conditioning, however, the CR remains sensitive to devaluation of the US after extended training, and hence it has been concluded that control of the CR by US representations

does not diminish after extended training [11-13]. To our knowledge, loss of sensitivity of the CR to US devaluation by extended Pavlovian training has not been reported in any animals.

In this study, we investigated whether extended Pavlovian training leads to a change of sensitivity of CR to devaluation of US in the cricket, *Gryllus bimaculatus*. Crickets are newly emerging animal models for the study of basic mechanisms of learning and memory, because they have excellent learning capabilities and it is relatively easy to apply various experimental manipulations such as pharmacological intervention, RNAi and genome editing [14-19]. We previously showed that 2 to 4 pairing trials to associate an odor with water, with inter-trial intervals (ITIs) of 5 min, are sufficient to lead to protein synthesis-dependent long-term memory that lasts for at least four days [20,21]. We refer to this training regimen with a single session of 4 trials with 5-min ITIs as standard training. For extended training, we applied three training sessions for one session per day on three consecutive days. We show that execution of the CR is sensitive to devaluation of the US after standard training but not after extended Pavlovian training in crickets. Moreover, we previously suggested that a class of octopamine (OA, invertebrate counterpart of noradrenaline) neurons conveys appetitive US signals in Pavlovian conditioning in crickets [14-16, 22-25] and that execution of the CR after standard training requires CS-induced activation of another class of OA neurons, which may convey US value signals that guide execution of the CR [16]. Here we show that control of the CR by OA neurons is lost after extended training. We thus propose that control of the CR by OA neurons decreases with extended training and this leads to the CR more independent of current value of the US.

2. Methods

(a) Animals

Adult male crickets, *Gryllus bimaculatus*, at 1 week after the imaginal molt were used. They were reared in a 12 h:12 h light:dark cycle at $27 \pm 2^\circ\text{C}$ and were fed a diet of insect pellets and water *ad libitum*. Three days before the start of the experiment, about 50 animals were individually placed in 100-ml glass beakers and deprived of drinking water to enhance their motivation to search for water.

(b) Classical olfactory conditioning

We used an elemental appetitive conditioning procedure in which an odor (CS) was paired with water (US) as described previously [20]. A syringe containing water was used for conditioning. A filter paper soaked with an odor essence was attached to the needle of the syringe. The filter paper was placed within 1 cm of the cricket's head for 3 sec so as to present an odor, and then a drop (ca. 10-13 μl) of water was presented to the mouth. After the conditioning trials, the air in the beaker was ventilated. The cricket was kept in a beaker until the next training or test. Crickets were subjected to 4 conditioning trials with 5-min ITIs for one day (4 trials x 1 day training) or for three consecutive days (4 trials x 3 days training).

For conditioning with sucrose solution, crickets were not given food for three days with water provided *ad libitum* and were then subjected to 4-trial training to associate an odor with 0.5 M sucrose solution with 5-min ITIs.

(c) Odor preference test

All groups of animals were subjected to preference tests between the odor used in training and an odor not used in training (control odor) before and after training [14,20]. The test apparatus consisted of two waiting chambers and a test chamber. On the floor of the test chamber, there were two holes that connected the chamber with two odor sources.

Each odor source consisted of a plastic container containing a filter paper soaked with 3 μl solution of odor essence, covered with a fine gauze net. Three containers were mounted on a rotatable holder and two of three odor sources could be located simultaneously just below the holes of the test chamber. Before the odor preference test, a cricket was transferred to the waiting chamber at the waiting position and left for about 4 min to become accustomed to the surroundings. The waiting chamber was then slid into the entrance position, and the door to the test chamber was opened. When the cricket entered the test chamber, the door was closed and the test was started. Two min later, the relative positions of the odor sources were changed by rotating the container holder. The preference test lasted for 4 min. If the total time of visits of an animal to either source was less than 10 sec, we considered that the animal was less motivated to visit odor sources, possibly due to a poor physical condition, and the data were rejected. Animals that fell into this category were less than 10% of all animals. After the test, the cricket was returned to a beaker. In the experiment in which the post-training test was performed 3 days after training, two drops (ca. 20-26 μl) of water were given every day until the day of the test to prevent excessive hydration.

We used two different pairs of odors for experiments. The first was peppermint odor and vanilla odor. Crickets prefer vanilla odor over peppermint odor [20] and thus we used peppermint odor as a CS to be associated with water, and vanilla odor was used as a control odor in a test to study relative preference between the CS and the control odor. Experiments with this odor pair led to a high learning score and allowed to obtain significant statistical results with relatively small number of samples [20] and thus are used for initial exploratory stages of this study. In experiments to examine whether results obtained with this odor pair can be generalized to another odor pair, apple odor and banana odor were used. Crickets prefer these two odors almost equally [14], and we used one of the two odors as a CS and the other odor as a control odor in a counterbalanced manner. Groups in which either apple odor or banana odor was used as

the CS showed similar learning scores [14], and thus data from the groups were pooled.

In the test, an odor source was considered to have been visited when the cricket probed the top net with its mouth. The time spent probing the top net of each odor source was recorded cumulatively. Relative odor preference of each animal was determined by using the preference index (PI), defined as $(t_r/(t_r+t_{nr})) \times 100$ (%), where t_r was the time spent exploring the odor associated with reward and t_{nr} was the time spent exploring the odor not associated with reward.

(d) Water satiation procedure

One day after completing the standard (4 trials x 1 day) training or extended (4 trials x 3 day) training, crickets were given water to their mouths by a syringe until they stopped consuming the water. They typically consumed 130-230 μ l of water. At 1 hour after the water satiation (or devaluation) trial, they were given a posttraining test. In parallel experiments, we observed that when tested 1 hour after the satiation procedure crickets typically consumed only 0-40 μ l of water, confirming that the satiation procedure was successful.

Drug injection

By using a 10 μ l syringe, saline solutions (each 3 μ l) containing drugs were injected into the heads of the crickets from a small hole made in the median ocellus. Epinastine, a highly effective antagonist of insect octopamine (OA) receptors (see discussion in [25]), was purchased from SIGMA (Tokyo, Japan). The drug was dissolved in cricket saline [21] (Matsumoto et al., 2003). The dose of epinastine was determined on the basis of previous studies [14,15].

(e) Experimental design and statistical analyses

We performed between-group and within-group comparisons of odor preferences to

examine if the two comparisons produce consistent results. Wilcoxon's signed-rank test (WCX test) was used to compare odor preferences before and after training of a given group (within-group comparison), and the Mann-Whitney test (M-W test) was used to compare odor preferences between the experimental and control groups, since the data often deviated from a normal distribution. Kruskal-Wallis test was used to compare odor preferences in the test prior to the training among groups used in this study. Sample sizes are determined by intrinsic variation of the data set. The post-hoc Holm's method was used to adjust the P value when comparing among three groups. All statistical results are described in supplementary table S1. P values of <0.05 were considered statistically significant. We used R (ver. 3.3.1) package for statistical analysis.

3. Results

(a) Effect of water satiation on execution of the conditioned response to water-associated odor after standard training.

We used two different odor pairs for conditioning experiments, namely, peppermint odor and vanilla odor pair and apple odor and banana odor pair. In initial stages of this study we used the former odor pair for experiments, while the latter odor pair was used for experiments to confirm results obtained in earlier stages (see Methods).

At first, two groups of crickets were deprived of water for three days and then subjected to 4-trial training (or 4 trials x 1 day training) to associate peppermint odor with water reward. Relative preference between peppermint odor (CS) and vanilla odor (control odor) was tested before and one day after training. Crickets in one group (water satiated or reward devalued group) were given water until they stopped drinking, and 1 hour after they were given post-training test. Separate experiments showed that crickets were sufficiently satiated with water at the time of the test (see Methods). Crickets in another group (control group) were not given water prior to the post-training test.

The control group exhibited significantly increased preference for the conditioned odor after training compared to that before training (figure 1a, for details of the statistical results, see supplementary table S1). We previously showed that this change is pairing-specific and not due to a non-associative effect caused by exposure to the odor or water [20]. On the other hand, water-satiated group did not exhibit significantly increased preference for the CS. Between-group comparison confirmed that the posttraining preference for the CS was significantly lower in water-satiated group than in the control group (figure 1a). This is in contrast to that Kruskal-Wallis test showed no significant differences in odor preference in the pre-training test among groups used in this study (supplementary table S1). We thus conclude that water satiation diminishes conditioned responses to water-associated odor. We obtained the same conclusion when we evaluated the data by the time spent visiting each odor (supplementary figure S1).

Is suppression of response to the CS by providing water until satiation specific to the response to the CS associated with water? To address this question, two groups of crickets were given no food for three days and then received 4-trial conditioning of peppermint odor with 0.5 M sucrose solution [24]. On the next day, crickets in one group were given water until satiation and those in the other group were not. The water satiated group exhibited a significant increase of the preference for the CS, the level of which was as high as that of the control group in which no water satiation procedure were performed (figure 1b). A parallel experiment with another group of crickets that received 4-trial conditioning of the same odor with water and were then given water until satiation reproduced the results in water-satiated group shown in figure 1a, namely, they exhibited no significant level of conditioned response to the CS (figure 1b). Thus, the effect of providing water until satiation is specific to the response to the CS associated with water US.

It could be argued that absence of increased preference for the water-associated odor in water-satiated group in figure 1a (and in supplementary figure S1) might possibly be due to that (1) the water satiation procedure changed the unconditioned preference for the odor, namely, reduction of preference for peppermint odor or (2) it led to extinction of memory, rather than inhibition of responses to the CS. The former possibility was ruled out by our observation that untrained crickets that had been given water until satiation exhibited no change of odor preference (supplementary figure S2). The later possibility is also unlikely since we observed in figure 1b that crickets that had received water-satiation procedure exhibited increased preferences for sucrose-associated peppermint odor. We thus conclude that devaluation of water US after standard (4 trials x 1 day) training leads to suppression of responses to water-associated CS. Moreover, the results of the experiment with extended training, which will be described next, serve as another evidence against the account based on non-specific effect of water satiation.

(b) Effect of reward devaluation on execution of the conditioned response after extended training.

Next, we performed reward devaluation experiments with 4 trials x 3 days training with water used as US. Peppermint odor was used as the CS and vanilla odor was used as the control odor. One group was subjected to the water satiation procedure prior to the post-test whereas the other group was not subjected to the procedure. Both the devalued and control groups exhibited significantly increased preference for the CS after training as compared to the pre-training test (figure 2a), indicating that responses to the water-associated CS (CRs) were not inhibited by water satiation after extended (4 trials x 3 days) training. This was in contrast to observations in figure 1a in which the CR was fully inhibited by water satiation after standard training. Contrasting effect of water satiation on CRs after standard and extended trainings confirms that inhibition of the CR by water satiation after standard training was not due to unspecific effects of water satiation on execution of the CR.

It could be argued that the loss of sensitivity of the CR to the US devaluation procedure after 4-trials x 3 days training shown in figure 2a might be due to passage of time for three days after the start of training, not due to increased amount of training. To examine this possibility, crickets in two groups were subjected to 4-trial training. Three days later, one group (devalued group) was given a test 1 hour after being given water until satiation and the other group (control group) was given a test without receiving water (figure 2b). The control group exhibited significantly increased preference for the CS after training compared to that before training but the devalued group did not. The results indicate that the passage of time is not the reason for the loss of sensitivity of the CR to water satiation procedure after extended training.

We next investigated whether the observed loss of sensitivity of the CR to water satiation procedure by extended training is reproducible in experiments with apple odor and banana odor, in which one of the odors was used as CS and the other was used as a

control odor in a counterbalanced manner. The results confirmed that the CR is initially sensitive to water satiation procedure after standard (4 trials x 1 day) training (figure 2c) but that the sensitivity is lost after extended (4 trials x 3 days) training (figure 2d). The results confirm the effect of extended training observed in figure 2a.

We next performed experiments to determine whether the total number of trials (12 trials) or the intervals between training sessions (1-day intervals) is more important to lead to insensitivity of the CR to US devaluation. We tested 12 trials x 1 day training and 6 trials x 2 days training, the number of trials being the same as that of 4 trials x 3 days training, using peppermint odor as CS and vanilla odor as the control odor. In the experiments with 12 trials x 1 day training, the control group exhibited significantly increased preference for the CS after training compared to that before training, but the devalued group did not (supplementary figure S3a), indicating that the CR is sensitive to reward devaluation. In experiments with 6 trials x 2 days training, both the control and devalued groups exhibited significantly increased preference for the CS after training compared to that before training (supplementary figure S3b), but between-group comparison showed that the preference for the CS after training in the devalued group was significantly lower than in the control group. This indicates that 6 trials x 2 days training is not sufficient to make the CR completely insensitive to US devaluation. We conclude that increased number of trials is not sufficient to lead to the loss of sensitivity of the CR to US devaluation. Rather, repetitions of trials with sufficient intervals are needed for this change.

(c) A neural model for the CR to lose sensitivity to reward devaluation by extended training.

How can the decrement of control of the CR by current value of the US with extended training be accounted for by changes in synaptic transmission in the neural circuitry mediating the learned behavior? To help answer this question, we propose a neural

circuit model (figure 3) by modifying our previously proposed model [16,22,23], which was assumed to represent the neural circuitry in lobes of the mushroom body, a secondary olfactory center that plays critical roles in olfactory learning [26,27]. The model was designed to account for the findings that administration of an octopamine (OA) receptor antagonist impairs both acquisition of conditioning and execution of CRs after standard Pavlovian training [14,15,18]. Our model assumes two classes of OA neurons, namely, OA1 neurons that govern learning process and OA2 neurons that govern execution of conditioned response [22,23], and in the model shown in Fig. 3 we focus on the roles of OA2 neurons. In the model, we assume that (1) presentation of the CS activates “CS” neurons and then “OA2” neurons, (2) activation of both types of neurons is needed to activate “CR” neurons that produce the CR after standard training (figure 3a) [16], (3) activation of “OA2” neurons is inhibited when animals are satiated with water and (4) the requirement of activation of “OA2” neurons for production of the CR is lost after extended training, being due either to enhancement of the efficacy of synaptic transmission from “CS” neurons and “CR” neurons in this circuitry by extended training, so that the requirement of activation of “OA2” neurons for activation of “CR” neurons is lost (figure 3b), or by the development of a novel neural circuitry that can produce the CR without activation of “OA2” neurons outside the area where initial learning occurs (figure 3c). In short, the model proposes that a type of OA neurons represents current value of the US and that CS-induced activation of the OA neurons controls the CR early in training, but the control is lost and hence the CR becomes independent of current value of the US with extended training.

(d) Effect of blockade of OA receptors on execution of the CR after standard training and extended training.

The model predicts that blockade of OA-ergic transmission impairs execution of the CR after standard training but not after extended training. We next investigated whether

this indeed occurs. At first, two groups were subjected to standard (4 trials x 1 day) training, and animals in these groups received injection of saline or saline containing epinastine, a potent antagonist of insect OA receptors [26], before the post-test (control and epinastine groups). Peppermint odor was used as CS and vanilla odor was used as the control odor. The control group exhibited a significantly increased preference for the CS after training compared to that before training, but the epinastine group did not (figure 4a). We previously showed that injection of epinastine leads to suppression of the CR but not impairment of memory retention [16]. In experiments with extended (4 trials x 3 days) training, on the other hand, both the control and epinastine groups exhibited a significantly increased preference for the CS after training compared to that before training (figure 4b). Thus, activation of OA neurons is needed for expression of the CR after standard training but not after extended training, being in accordance with our model.

To confirm the reproducibility of the results shown in figure 5a and b, similar experiments were performed with either apple odor or banana odor being used as CS and the other being used as the control odor. In experiments with standard training, the control (saline-injected) group exhibited a significantly increased preference for the CS after training, but the epinastine group did not (figure 4c). On the other hand, in experiments with extended training, both the control group and epinastine group exhibited a significantly increased preference for the CS after training (figure 4d). The results confirm that normal synaptic transmission from OA neurons is needed for execution of the CR after standard training but not after extended training, regardless of the kinds of odors used in experiments.

Next, we performed experiments to determine whether the amount of training required to make the CR insensitive to OA receptor blockade matches that required to make the CR insensitive to reward devaluation. In experiments with 12 trials x 1 day training, the saline-injected control group exhibited a significantly increased preference

for the CS after training compared to that before training, but the epinastine group did not (supplementary figure S4a). In experiments with 6 trials x 2 days training, the saline-injected control group exhibited a significantly increased preference for the CS after training (supplementary figure S4b), but the epinastine group did not. Between-group comparison, however, showed that the post-training preference for the CS in the epinastine group did not significantly differ from that in the control group. Thus, 6 trials x 2 days training is not sufficient to make the CR completely insensitive to blockade of OA receptors. We conclude that the amount of training to make the CR insensitive to blockade of OA receptors is, in general, in accordance with that to make the CR insensitive to US devaluation, although the match might not be perfect.

4. Discussion

This study demonstrated a novel feature of Pavlovian conditioning, namely, execution of the CR is initially sensitive to devaluation of US (reward) but becomes insensitive to reward devaluation after extended training in crickets. Many studies in instrumental conditioning in mammals (mostly in rats) suggest that learned actions are initially goal-directed and are sensitive to devaluation of the outcomes but that they become more insensitive to outcome devaluation after extended training [1-3], although it should be cautioned that such change does not occur in some types of training [1,4,5]. On the other hand, no evidence showing a decrement of sensitivity of the CR to US devaluation by extended Pavlovian conditioning has been reported in mammals [11-13]. To our knowledge, this is the first study showing decrement of the control of the CR by current value of the US with extended Pavlovian training in any species of animals.

There is evidence, however, that contents of learning change with extended Pavlovian training in rats [11-13]. For example, after minimal training to associate an auditory or visual CS with a food US, consumption of food was reduced when the CS was paired with toxin (LiCl), indicating the CS-evoked representation of the food US entered into association with the toxin [11,12]. After extended training, however, the pairing of the CS with toxin did not reduce consumption of food, due possibly to the reduction in the effectiveness of the CS to activate sensory aspects of representation of the US after extended training. On the other hand, conditioned responding to the CS was reduced when the US was paired with toxin after minimal and extended trainings, due possibly to that the effectiveness of the CS to activate motivational aspects of US representation does not change with the extension of training [11,12]. Therefore, the effect of extended training to reduce the contents of learning in rats is much subtler than that in crickets.

It should also be cautioned that we used satiation procedure for reward devaluation in crickets, while taste aversion learning has been used in most studies in rats [1,11,12]. Future studies are needed to investigate whether different effect of reward devaluation

on CR after extended Pavlovian training in rats and crickets might have been due to different procedures of reward devaluation.

We suggest that the loss of sensitivity of CR to reward devaluation by extended training is due to loss of control of the CR by OA neurons that may mediate value signals of the US. Interestingly, it has been reported in Pavlovian conditioning in rats that blockade of dopaminergic (DA-ergic) transmission impairs execution of CR after minimal training but the sensitivity of the CR to blockade of DA-ergic transmission decreases with extended training [28,29]. It is not known, however, whether the decrease of DA-dependency of CR accompanies a decrease of sensitivity to US devaluation procedure in rats.

We propose an updated conceptual definition of habit formation to match the fact that the loss of sensitivity of learned behavior to reward devaluation can occur in both instrumental conditioning and Pavlovian conditioning. Theorists argue that learned action early in instrumental training in mammals often depends on action-outcome (A-O) association but that the action becomes dependent more on stimulus-response (S-R) association with the progress of training [1-3, 30]. In Pavlovian conditioning in crickets, it can be argued from our model that execution of the CR requires activation of both stimulus-stimulus (S-S) association and S-R association early in training but that the CR becomes dependent solely on the S-R association after extended training (see legend of figure 4). Thus, habit formation can be defined as learned behavior becoming dependent more on S-R association by extended training, regardless of the associative learning being instrumental or Pavlovian. In rats, distinction between S-R systems and S-S systems have been well documented in Pavlovian conditioning [31] and related tasks [2,32]. Both the systems form from earlier stage of training and which of the two systems has more influence on behavior depends on the type of training [33]. Since the distinction between S-S and S-R systems (and between A-O and S-R systems) matches the distinction between model-based and model-free systems in reinforcement learning theories [6],

habit formation can also be considered as learned behavior becoming dependent more on a model-free system after extended training.

In rats, it has been reported that reward devaluation reduces CRs after first- but not second-order conditioning, and argued that S-R association but not S-S association underlies CRs after second-order conditioning [34]. In crickets, we have proposed a model in which both S-S and S-R associations underlie CRs after second-order conditioning [16], and the model predicts that CRs after second-order conditioning are sensitive to reward devaluation. This prediction needs to be tested.

What is the possible functional significance of the changes of associative substrates underlying the CR with the progress of Pavlovian training? In instrumental conditioning, it has been argued that advantages of “action-to-habit” transition include (1) making higher order cognitive centers free from routine tasks and (2) allowing more rapid responses [7]. Both might be applicable to Pavlovian conditioning, but the former might be particularly important for insects because insects have very small brains and thus neuronal resources that can be allocated for new learning would be limited.

In rodents, there is evidence showing that the shift from initial goal-directed action to habitual response by extended instrumental training accompanies a shift in the neural substrates subserving memory storage and execution of learned actions [7,9,10]. Therefore, our next step should be to study the neural circuitry that mediates changes in the CR with extended Pavlovian training, and the model proposed in figure 4 should provide a basis for this study. There is evidence in some insects that olfactory learning initially occurs in the mushroom body [26,27,35], and investigation is needed to determine whether the mushroom body or another area of the central olfactory pathway (figure 4b, c), such as the antennal lobe (primary olfactory center), mediates the CR after extended Pavlovian training. Studies with microinjections of epinastine into specific brain regions, as well as immunohistochemical and electrophysiological characterization of OA neurons projecting to the mushroom body or the antennal lobe, will help to clarify

this point.

In fruit-flies, it has been reported that memory formed by Pavlovian conditioning by a small number of training sessions was generalized or transferred to a new behavior, but the generalization decreased with increase in the training sessions and, moreover, evidence has been provided to show that the mushroom body is needed for this generalization [36]. That study demonstrated that a change in neural substrates subserving learned behavior occurs by extended training in fruit-flies. Investigation is needed to determine whether this training accompanies a change in sensitivity to US devaluation procedure.

Psychologists argue that not only actions but also the way of thinking in humans (the way in which humans deal with sensory or social experiences) may become more stereotyped or automatic by repetitive sensory or social experiences, and this may lead to psychological disorder [37,38]. Thus, it is of great theoretical interest to understand whether repetitive experiences of a given stimulus-stimulus contingency (correction) could give rise to a habit-like inflexible cognition or response in humans and other animals. In the current study, we offered the first demonstration of such a case in crickets.

DATA accessibility. A summary of the data for each experiment is provided in a PDF file as the electronic supplementary material.

Author contributions. MM designed and SH, AS, RA, KT and YM performed behavioral and pharmacological experiments, SH and KT performed statistical analysis, and MM wrote the manuscript with the help of other authors. All authors gave final approval for publication.

Acknowledgements. We thank Yutaka Kosaki for helpful comments on the manuscript.

Competing interests. The authors have no competing interests.

Funding. This study was supported by Grants-in-Aid for Scientific Research from the Ministry of Education, Science, Culture, Sports and Technology of Japan (No. 16H04814 and 16K18586 to MM) and that to JSPS fellows (No. 15J01414 to KT).

References

1. Dickinson A. 1985 Action and habits: the development of behavioral autonomy. *Philos. Trans. R. Soc. Lond. B.* **308**, 67-78.
2. Yin HH, Knowlton BJ. 2006 The role of the basal ganglia in habit formation. *Nature Rev. Neurosci.* **7**, 464-476.24.
3. Smith KS, Graybiel AM. 2014 Investigating habits: strategies, technologies and models. *Front. Behav. Neurosci.* **8**, 39.
4. Kosaki Y, Dickinson A. 2010 Choice and contingency in the development of behavioral autonomy during instrumental conditioning. *J. Exp. Psychol. Anim. Behav. Process.* **36**, 334-342.
5. DeRusso AL, Fan D, Gupta J, Shelest O, Costa RM, Yin HH. 2010 Instrumental uncertainty as a determinant of behavior under interval schedules of reinforcement. *Front. Integr. Neurosci.* **4**, 17.
6. Dolan RJ, Dayan P. 2013 Goals and habits in the brain. *Neuron* **80**, 312-325.
7. Wood W, Runger D. 2016 Psychology of Habit. *Annu. Rev. Psychol.* **67**, 289-314.
8. Everitt BJ, Robbins TW. 2016 Drug addiction: Updating actions to habits to compulsions ten years on. *Annu. Rev. Psychol.* **6**, 23-50.
9. O'Hare JK, Ade KK, Sukharnikova T, Van Hooser SD, Palmeri ML, Yin HH, Calakos N. 2016 Pathway-specific striatal substrates for habitual behavior. *Neuron* **89**, 472-479.
10. Gremel CM, Chancey JH, Atwood BK, Luo G, Neve R, Ramakrishnan C, Deisseroth K, Lovinger DM, Costa RM. 2016 Endocannabinoid modulation of orbitostriatal circuits gates habit formation. *Neuron* **90**, 1312-1324.
11. Holland P. 1998 Amount of training affects associatively-activated event representation. *Neuropharmacol* **37**, 461-469.
12. Holland PC 2005 Amount of training effects in representation-mediated food aversion learning: no evidence of a role for associability changes. *Learn. Behav.* **33**, 464-478.
13. Holland PC, Lasseter H, Agrawal I. 2008 Amount of training and cue-evoked

- taste-reactivity responding in reinforcer devaluation. *J. Exp. Psychol.* **34**, 119-132.
14. Unoki S, Matsumoto Y, Mizunami M. 2005 Participation of octopaminergic reward system and dopaminergic punishment system in insect olfactory learning revealed by pharmacological study. *Eur. J. Neurosci.* **22**, 1409-1416.
 15. Unoki S, Matsumoto Y, Mizunami M. 2006 Roles of octopaminergic and dopaminergic neurons in mediating reward and punishment signals in insect visual learning. *Eur. J. Neurosci.* **24**, 2031-2038.
 16. Mizunami M, Unoki S, Mori Y, Hirashima D, Hatano A, Matsumoto Y. 2009 Roles of octopaminergic and dopaminergic neurons in appetitive and aversive memory recall in an insect. *BMC Biol* **7**, 46.
 17. Mizunami M, Terao K, Alvarez B. 2018 Application of a prediction error theory to Pavlovian conditioning in an insect. *Front. Psychol.* **9**, 1272.
 18. Mizunami M, Matsumoto Y. 2017 Roles of octopamine and dopamine neurons for mediating appetitive and aversive signals in Pavlovian conditioning in crickets. *Front. Physiol.* **8**, 1027.
 19. Matsumoto Y, Matsumoto CS, Mizunami M. 2018 Signaling pathways for long-term memory formation in the cricket. *Front. Psychol.* **9**, 1014.
 20. Matsumoto Y, Mizunami M. 2002 Temporal determinants of long-term retention of olfactory memory in the cricket *Gryllus bimaculatus*. *J. Exp. Biol.* **205**, 1429-1437.
 21. Matsumoto Y, Noji S, Mizunami M. 2003 Time course of protein synthesis-dependent phase of olfactory memory in the cricket *Gryllus bimaculatus*. *Zool. Sci.* **20**, 409-416.
 22. Terao K, Matsumoto Y, Mizunami M. 2015 Critical evidence for the prediction error theory in associative learning. *Sci. Rep.* **5**, 8929.
 23. Terao K, Mizunami M. 2017 Roles of dopamine neurons in mediating the prediction error in aversive learning in insects. *Sci. Rep.* **7**, 14694.
 24. Awata H, Watanabe T, Hamanaka Y, Mito T, Noji S, Mizunami M. 2015 Knockout crickets for the study of learning and memory: Dopamine receptor Dop1 mediates

- aversive but not appetitive reinforcement in crickets. *Sci. Rep.* **5**, 15885.
25. Awata H, Wakuda R, Ishimaru Y, Matsuoka Y, Terao K, Katata S, Matsumoto Y, Hamanaka Y, Noji S, Mito T, Mizunami M. 2016 Roles of OA1 octopamine receptor and Dop1 dopamine receptor in mediating appetitive and aversive reinforcement revealed by RNAi studies. *Sci. Rep.* **6**, 29696.
 26. Liu C, Plaçais PY, Yamagata N, Pfeiffer BD, Aso Y, Friedrich AB, Siwanowicz I, Rubin GM, Preat T, Tanimoto H. 2012 A subset of dopamine neurons signals reward for odour memory in *Drosophila*. *Nature* **488**, 512-516.
 27. Giurfa M. 2013 Cognition with few neurons: higher-order learning in insects. *Trends Neurosci.* **36**, 285-294.
 28. Choi WY, Balsam PD, Horvitz JC. 2005 Extended habit training reduces dopamine mediation of appetitive response expression. *J. Neurosci.* **25**, 6729-6733.
 29. Clark JJ, Collins AL, Sanford CA, Phillips PE. 2013 Dopamine encoding of Pavlovian incentive stimuli diminishes with extended training. *J. Neurosci.* **33**, 3526-3532.
 30. Balleine BW, Dickinson A. 1998 Goal-directed instrumental action: contingency and incentive learning and their cortical substrates. *Neuropharmacol.* **37**, 407-419.
 31. Holland PC. 2008 Cognitive versus stimulus-response theories of learning. *Learn Behav* **36**, 227-241.
 32. Perks SM, Clifton PG. 1997 Reinforcer revaluation and conditioned place preference. *Physiol. Behav.* **61**, 1-5.
 33. Clark JJ, Hollon NG, Phillips PE. 2012 Pavlovian valuation systems in learning and decision making. *Curr. Opin. Neurobiol.* **22**, 1054-1061.
 34. Holland PC, Rescorla RA. 1975 The effect of two ways of devaluing the unconditioned stimulus after first- and second-order appetitive conditioning. *J. Exp. Psychol. Anim. Behav. Process.* **1**, 355-363.
 35. Menzel R. 2012 The honeybee as a model for understanding the basis of cognition. *Nat. Rev. Neurosci.* **13**, 758-768.

36. Brembs B. 2009 Mushroom bodies regulate habit formation in *Drosophila*. *Curr. Biol.* **19**, 1351-1355.
37. Verplanken B, Fisher N. 2014 Habitual worrying and benefits of mindfulness. *Mindfulness* **5**, 566-573.
38. Gardner B. 2015 A review and analysis of the use of 'habit' in understanding, predicting and influencing health-related behaviour. *Health Psychol. Rev.* **9**, 277-295.

Figure legends

Figure 1. Effects of water satiation on conditioned responses to water-associated odor after standard training. (a) Two groups were each subjected to a pre-training test and then 4-trial (4 trials x 1 day) training. On the next day, crickets in one of the two groups were given water at 1 hour prior to the post-training test until they stopped drinking (water satiated group) and those in the other group were not given water (control group). (b) The effect of giving water until satiation on responses to an odor CS after conditioning with sucrose US compared to that after conditioning with water US, was evaluated by comparing among three groups. Two of the three groups were each subjected to a pretest and then 4-trial training with 0.5 M sucrose reward. On the next day, crickets in one group were given water until satiation at 1 hour prior to the posttest and those in another group were subjected to the test without giving water. The last group was subjected to a pretest and then 4-trial training with water reward. On the next day, crickets in this group were given water at 1 hour prior to the posttest. The time schedules of experiments are shown above the graphs. Peppermint odor was used as CS and vanilla odor was used as the control odor (indicated as p/v). Relative preference for the rewarded odor, measured as the preference index (PI), before training (white box) and that 1 day after training (gray box) are shown as box plots. The results of statistical comparison before and after conditioning (WCX test) and between groups (M-W test) are shown as asterisks (* * $p < 0.01$; * * * $p < 0.001$; NS $p > 0.05$, with the p values adjusted by Holm's method).

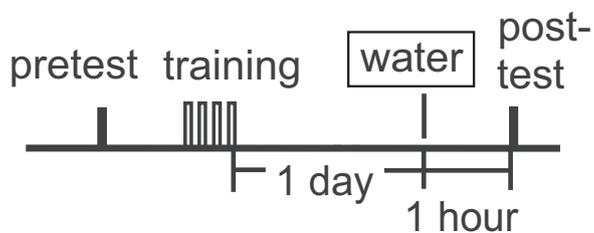
Figure 2. Effects of reward devaluation on execution of the CR after standard training and extended training. (a) Two groups were each subjected to a pretest and then 4 trials x 3 days training. On the next day, crickets in one of the two groups were given water at 1 hour prior to the posttest until they stopped drinking (devalued group) and those in the other group were not given water (control group). (b) A test was performed to determine possible effects of passage of time for three days after 4-trial training on sensitivity of the

CR to water devaluation. Two groups were each subjected to a pretest and then 4-trial training. Three days later, crickets in one group were given a sufficient amount of water prior to the posttest and those in the other group were not. The time schedules of experiments are shown above the graphs. Peppermint odor was used as CS and vanilla odor was used as the control odor (indicated as p/v). (*c,d*) The experiments for which results are shown in Figs. 1A and 2A were replicated with either apple odor or banana odor being CS and the other odor being the control odor (indicated as a:b). Relative preference for the rewarded odor, measured as the preference index (PI), before training (white box) and that 1 day after training (gray box) are shown as box plots. The results of statistical comparison before and after conditioning (WCX test) and between groups (M-W test) are shown as asterisks (** $p < 0.01$; *** $p < 0.001$; NS $p > 0.05$, with the p values adjusted by Holm's method).

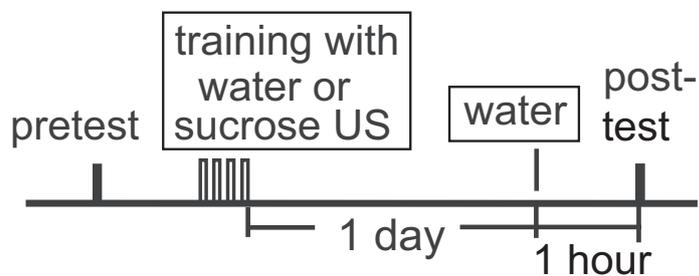
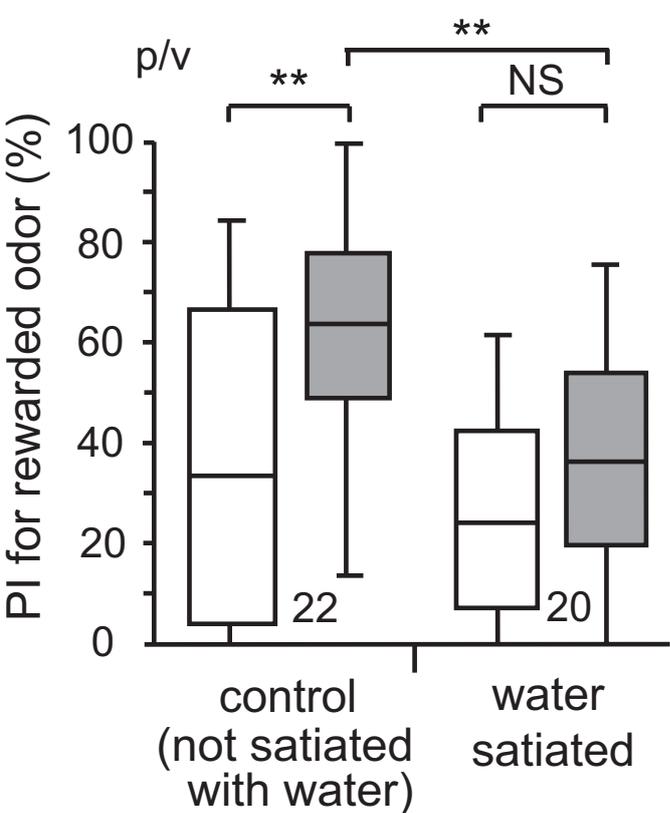
Figure 3. A model to account for the finding that the CR is sensitive to reward devaluation after standard training but not after extended training. The model was modified from our previous model proposed to account for the finding that activation of OA neurons is needed for execution of the CR after standard training [22,23]. Our original model contained two types of OA neurons, “OA1” neurons that participate in learning process and “OA2” neurons that participate in execution of CR, and in this model we focus on “OA2” neurons, which we assume to mediate current value of the US. (*a*) The basic assumptions of the model are that (1) it represents neural circuits in the lobe of the mushroom body (MB): “CS” neurons that encode CS and “CR” neurons for which activation leads to the CR represent Kenyon cells and output neurons of the MB lobe, respectively, (2) two kinds of memory traces are formed by standard training, one being synaptic connection from “CS” to “OA2” neurons, which represents the CS-US (S-S) connection, and the other being synaptic connection from “CS” neurons to “CR” neurons, which represents the CS-CR (S-R) connection, (3) simultaneous activation of “CS”

neurons and “OA2” neurons is needed for activating “CR” neurons (AND gate) and for producing the CR and (4) activation of “OA2” neurons is inhibited when the animal has been satiated with water. (b,c) The CR becomes insensitive to reward devaluation by extended training due to strengthening of the CS-CR connection and resulting elimination of the requirement of activation of “OA2” neurons for the CR (b) or by development of a novel neural circuitry that produces the CR without activation of “OA2” neurons independently of the neural circuitry where learning initially occurs (c).

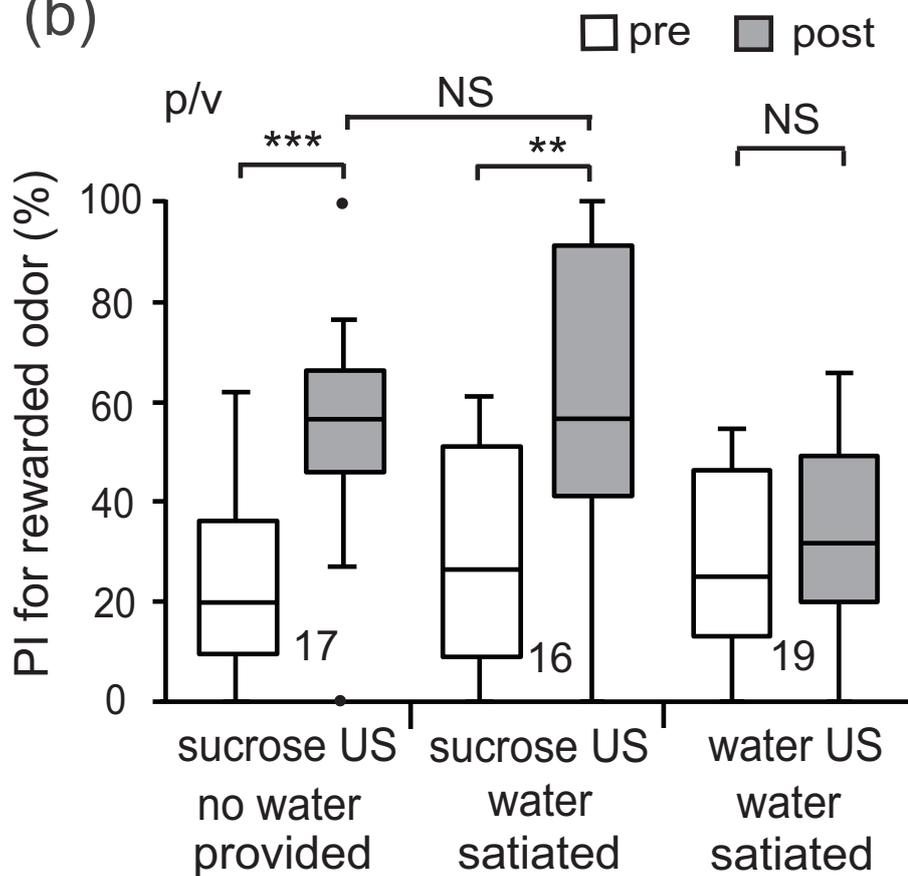
Figure 4. Effects of epinastine on execution of the CR after standard training and extended training. (a,b) Groups of animals were each subjected to a pretest and then 4 trials x 1 day training (a) or 4 trials x 3 days training (b). On the next day, crickets in each group were injected with 3 μ l of saline or saline containing 2 μ M epinastine at 30 min before the test. Peppermint odor was used as CS and vanilla odor was used as the control odor. (c,d) The experiments were replicated with either apple odor or banana odor being used as CS and the other being used as the control odor (indicated as a:b). Relative preferences for the rewarded odor (PI) before (white box) and 1 day after training (gray box) are shown as box plots. The results of statistical comparison before and after conditioning (WCX test) and between groups (M-W test) are shown as asterisks (* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$; NS $p > 0.05$, with the p values adjusted by Holm’s method).

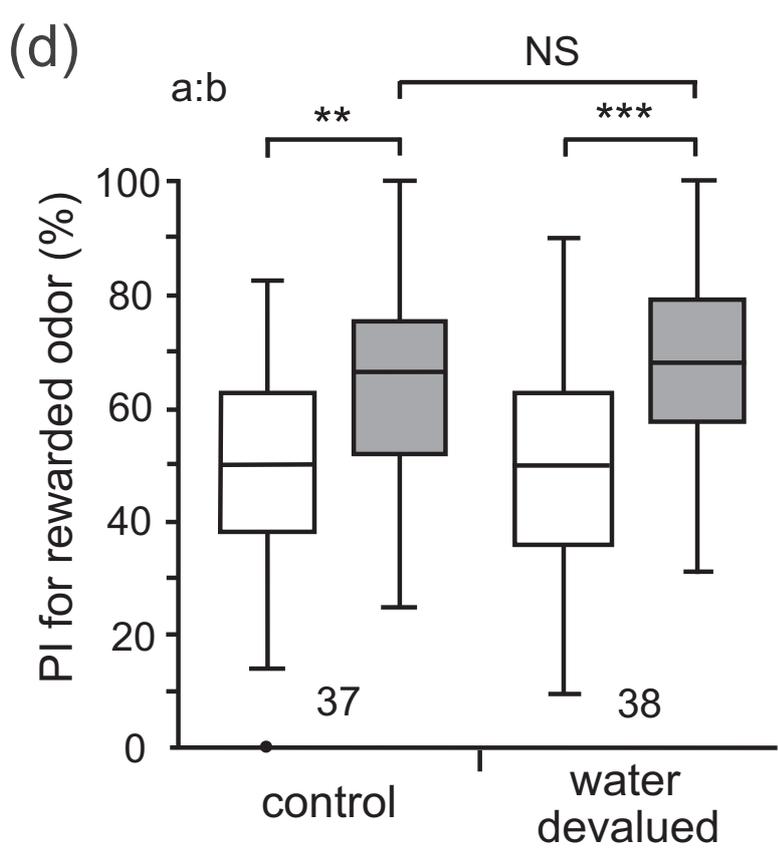
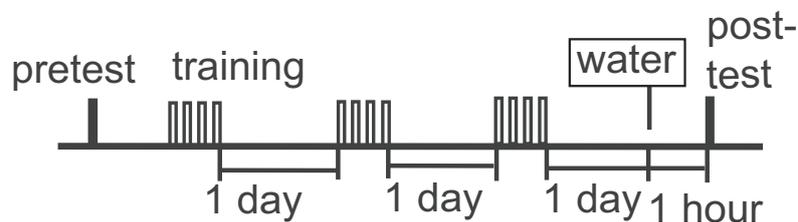
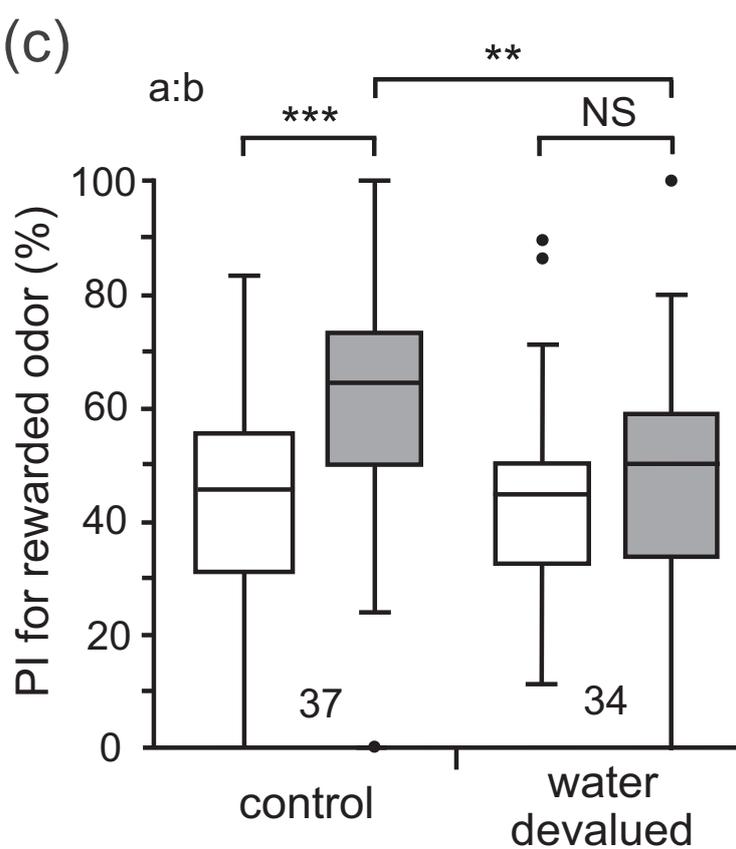
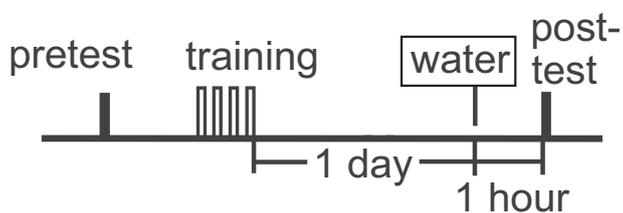
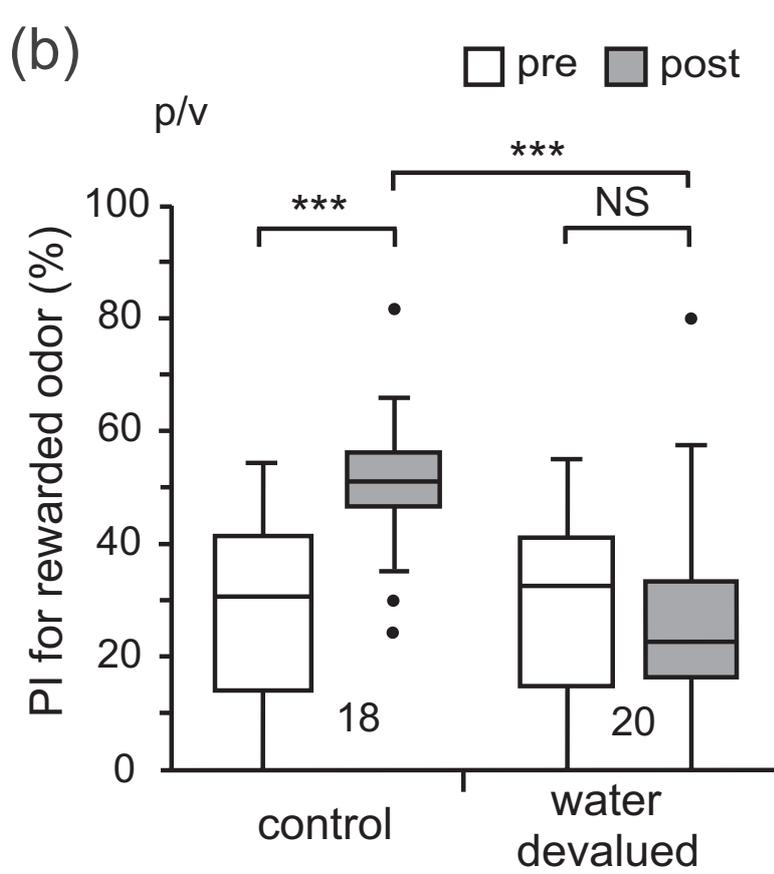
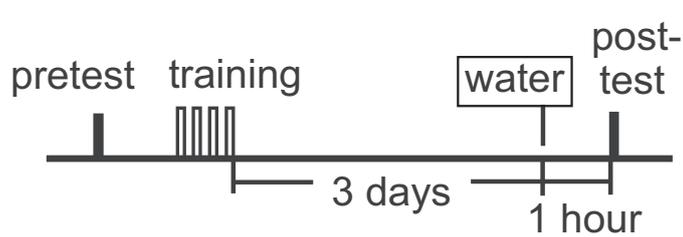
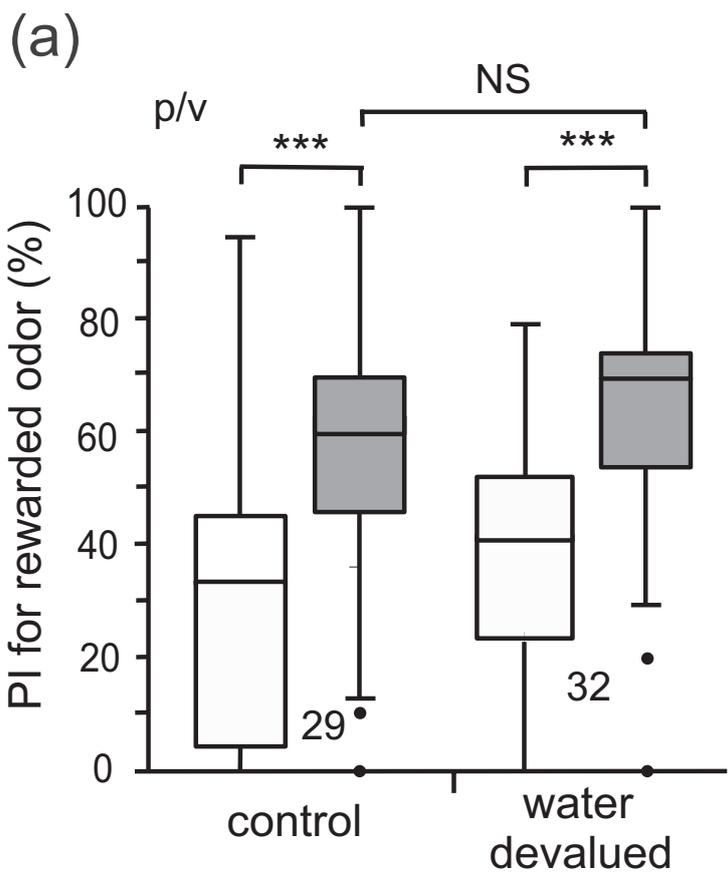
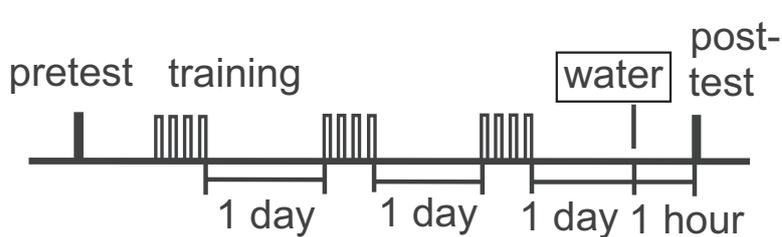


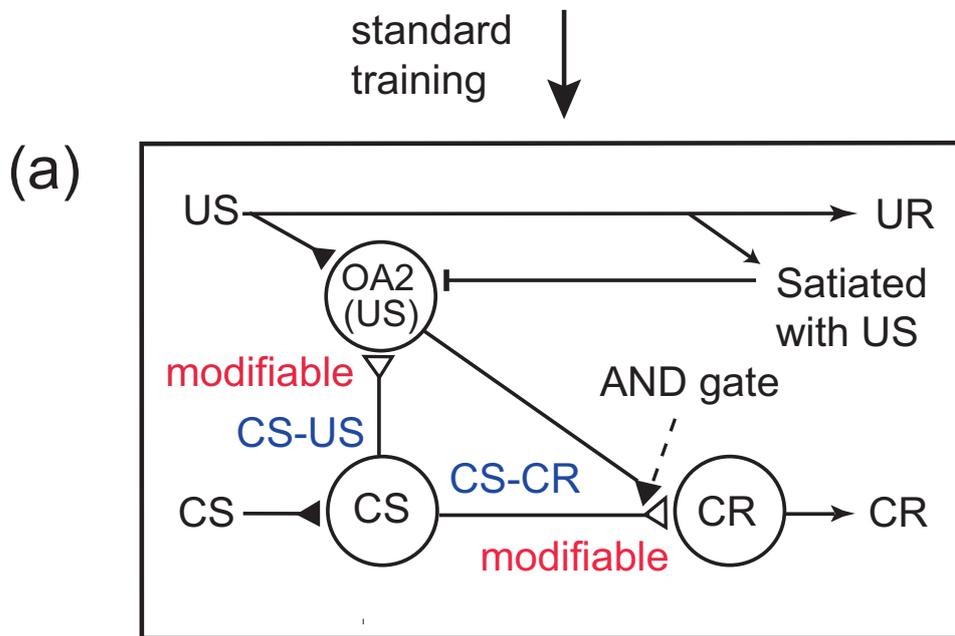
(a)



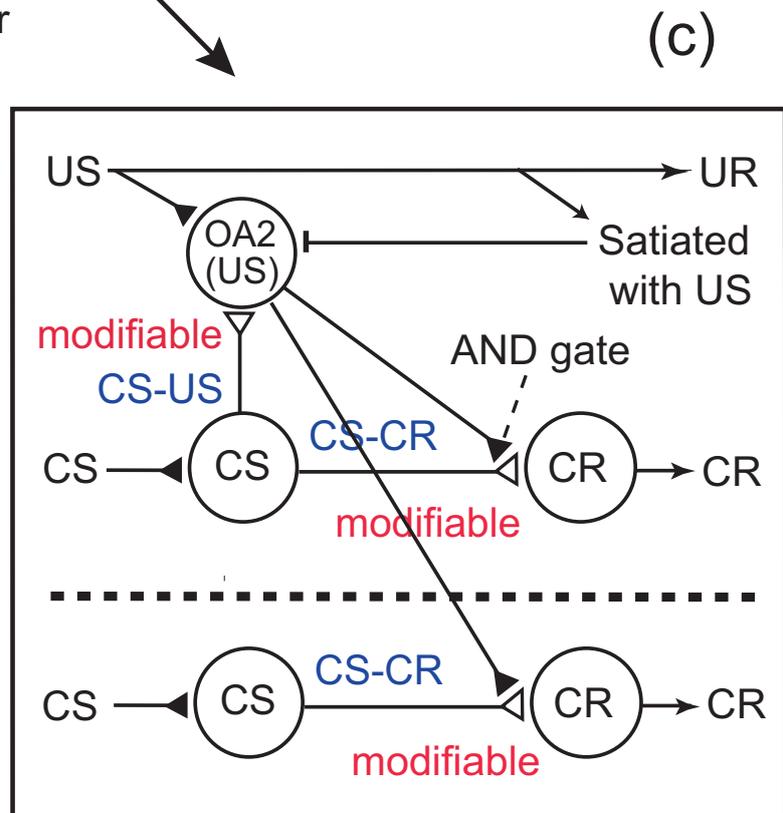
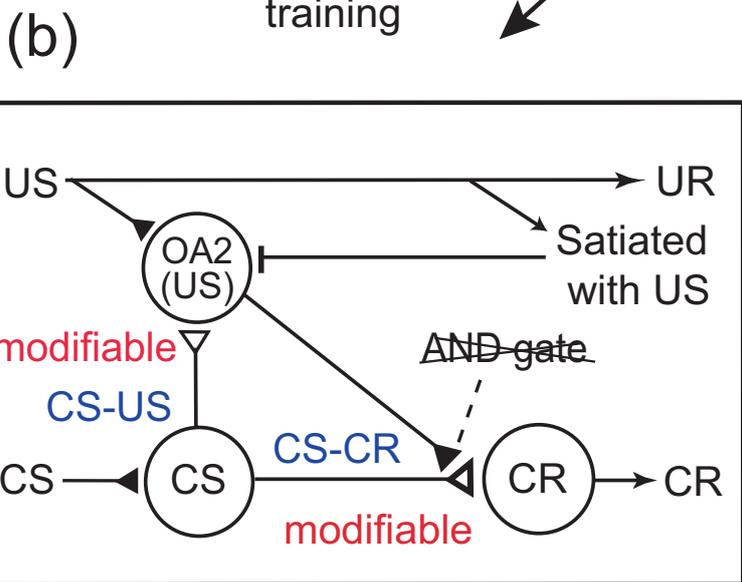
(b)

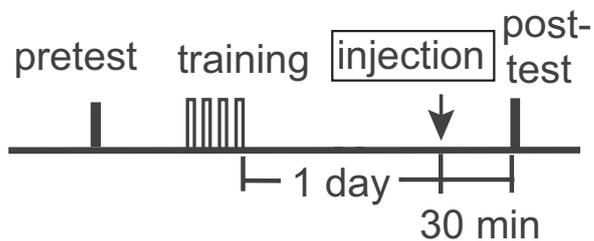




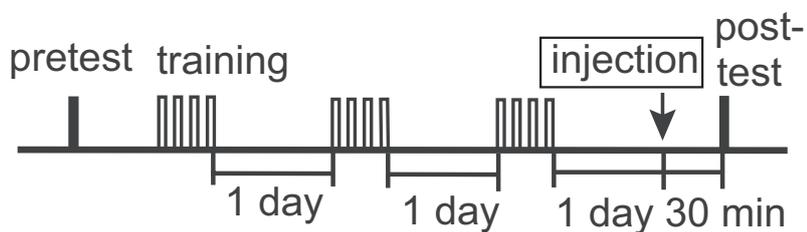
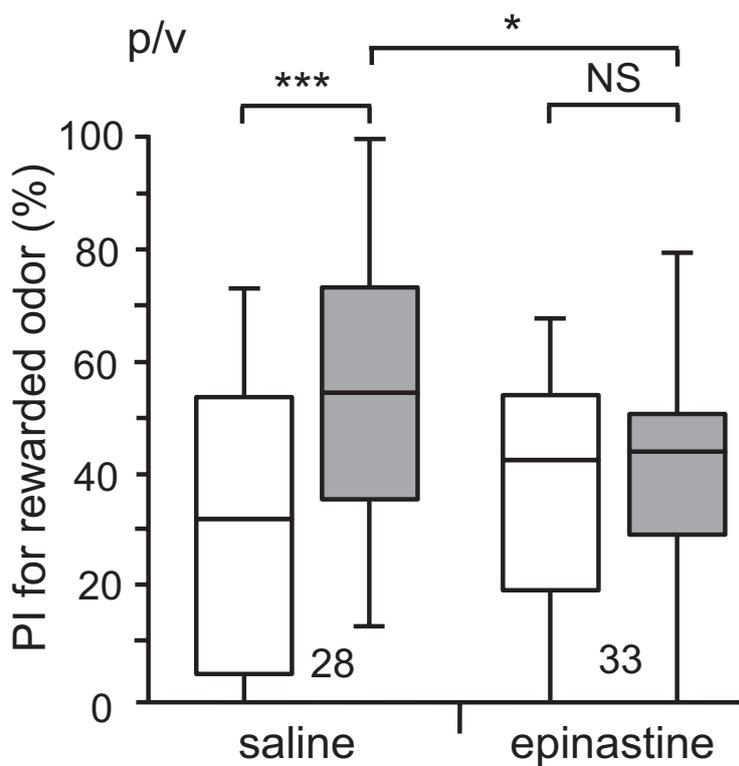


extended training ↙ or ↘

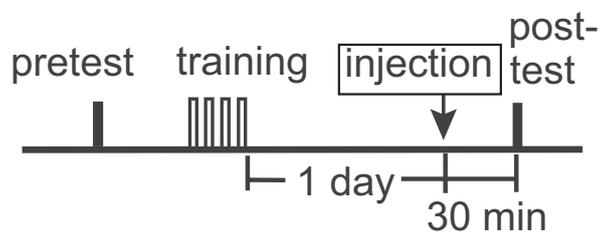
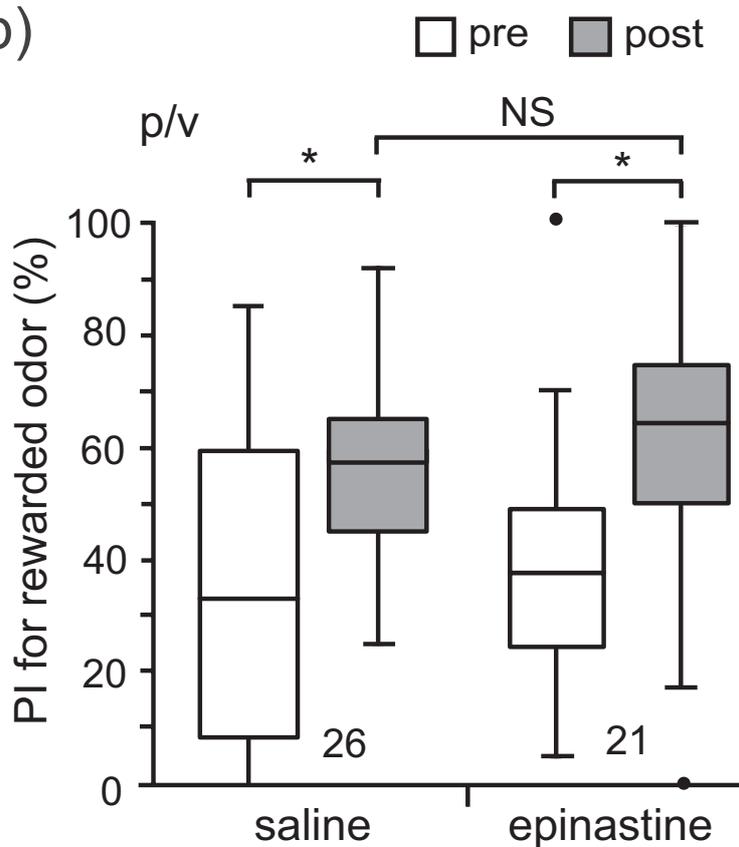




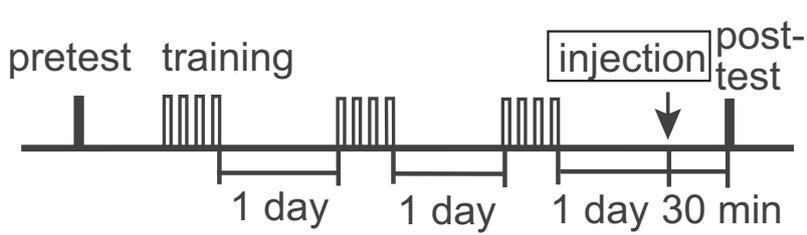
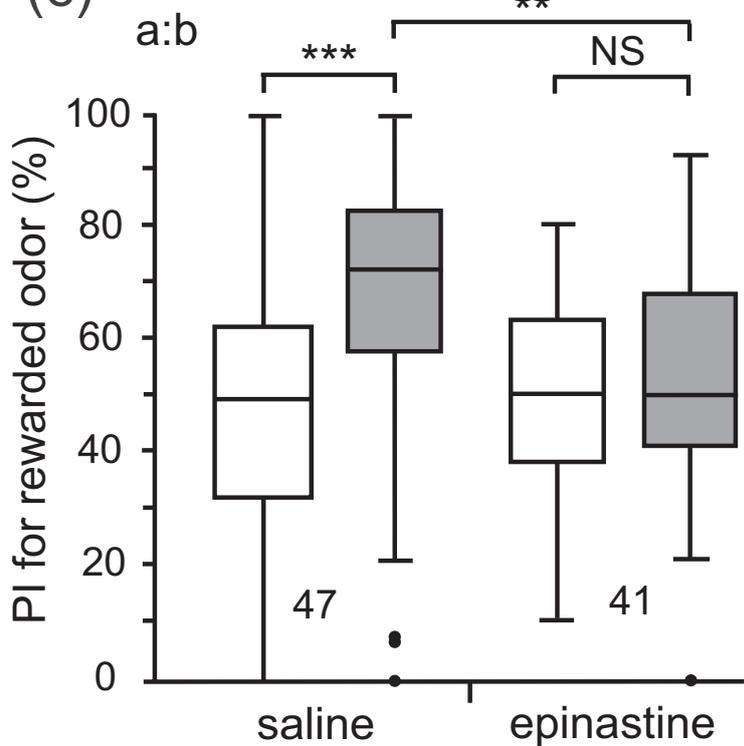
(a)



(b)



(c)



(d)

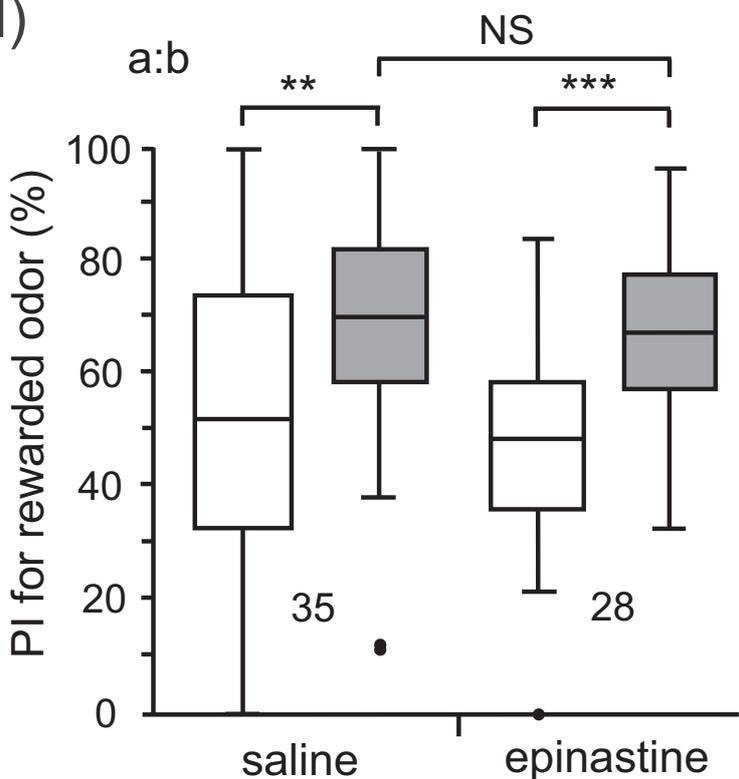


Table S1. Statistical results.

Groups	n	Test	H or W or U	P	Fig.
p/v, before (17 groups)		K-W	H=18.4 (df=19)	0.50	-
a:b, before (12 groups)		K-W	H=15.0 (df=15)	0.45	-
4trials x 1d, p/v; control: before vs after training	22	WCX	W=31	0.0033	1a
devalued: before vs after training	20	WCX	W=69	0.189	1a
After training: control vs devalued		M-W	U=346	0.0031	1a
4trials x 1d, p/v; sucrose US, no water: before vs after training	17	WCX	W=18	0.00038	1b
sucrose US, water: before vs after training	16	WCX	W=10	0.0049	1b
water US, water: before vs after training	19	WCX	W=66	0.41	1b
sucrose US, after training: no water vs water		M-W	U=161	0.38	1b
4trials x 3d, p/v; control: before vs after training	29	WCX	W=60	0.0036	2a
devalued: before vs after training	32	WCX	W=78	0.00013	2a
after training: control vs devalued		M-W	U=573.5	0.12	2a
4trials x 1d, p/v; control: before vs 3d after training	18	WCX	W=11	0.00084	2b
devalued: before vs 3d after training	20	WCX	W=123	0.52	2b
3d after training: control vs devalued		M-W	U=313	0.00032	2b
4trials x 1d, a:b; control: before vs after training	37	WCX	W=277	0.000067	2c
devalued: before vs after training	34	WCX	W=88	0.96	2c
After training: control vs devalued		M-W	U=368.5	0.0055	2c
4trials x 3d, a:b; control: before vs after training	37	WCX	W=142	0.0023	2d
devalued: before vs after training	38	WCX	W=130	0.00084	2d
after training: control vs devalued		M-W	U=785.5	0.38	2d
4trials x 1d, p/v; saline: before vs after training	33	WCX	W=72	0.00023	4a
epinastine: before vs after training	28	WCX	W=169	0.45	4a
after training: saline vs epinastine		M-W	U=619	0.047	4a
4trial x 3d, p/v; saline: before vs after training	26	WCX	W=95	0.026	4b
epinastine: before vs after training	21	WCX	W=117	0.029	4b
after training: saline vs epinastine		M-W	U=313.5	0.39	4b
4trials x 1d,a:b; saline: before vs after training	47	WCX	W=205	0.00022	4c
epinastine: before vs after training	41	WCX	W=351	0.43	4c
after training: saline vs epinastine		M-W	U=1351	0.0024	4c

Groups	n	Test	W or U	P	Fig.
4trials x 3d,a:b; saline: before vs after training	35	WCX	W=145	0.0090	4d
epinastine: before vs after training	28	WCX	W=52	0.00082	4d
after training: saline vs epinastine		M-W	U=525	0.73	4d
4trialx 1d, time to visit peppermint; control: before vs after training	24	WCX	W=40	0.0035	S1a
devalued: before vs after training	24	WCX	W=133.5	0.90	S1a
After training: control vs devalued		M-W	U=90.5	0.00014	S1a
4trials x 1d, time to visit vanilla; control: before vs after training	24	WCX	W=176	0.47	S1b
devalued: before vs after training	24	WCX	W=216.5	0.18	S1b
After training: control vs devalued		M-W	U=213.5	0.25	S1b
4trials x 1d, total time to visit odors; control: before vs after training	24	WCX	W=74	0.11	S1c
devalued: before vs after training	24	WCX	W=209	0.094	S1c
After training: control vs devalued		M-W	U=120.5	0.0017	S1c
no training p/v: before vs after water devaluation	19	WCX	W=116	0.19	S2
12 trials x 1d, p/v; control: before vs after training	22	WCX	W=3	0.0000072	S3a
devalued: before vs after training	25	WCX	W=175	0.48	S3a
after training: control vs devalued		M-W	U=438	0.0010	S3a
6trials x 2d, p/v; control: before vs after training	22	WCX	W=14	0.00016	S3b
devalued: before vs after training	24	WCX	W=62.5	0.023	S3b
after training: control vs devalued		M-W	U=388	0.013	S3b
12trials x 1d, a:b; saline: before vs after training	21	WCX	W=43	0.037	S4a
epinastine: before vs after training	25	WCX	W=136	0.48	S4a
after training: saline vs epinastine		M-W	U=364	0.026	S4a
6trials x 2d, a:b; saline: before vs after training	21	WCX	W=37	0.015	S4b
epinastine: before vs after training	17	WCX	W=48	0.63	S4b
after training: saline vs epinastine		M-W	U=212	0.33	S4b

p/v: peppermint odor was used as CS and vanilla odoe was used as control odor: a:b: either of the apple odor or banana odor was used as CS and the other as the control odor in a counterbalanced manner. K-W: Kruskal-Wallis test. WCX: Wilcoxon's test. M-W: Mann-Whitney test. P values were adjusted by Holm's method in the case of multiple comparisons. Experiments with the K-W test are described in the third paragraph of the Results section.

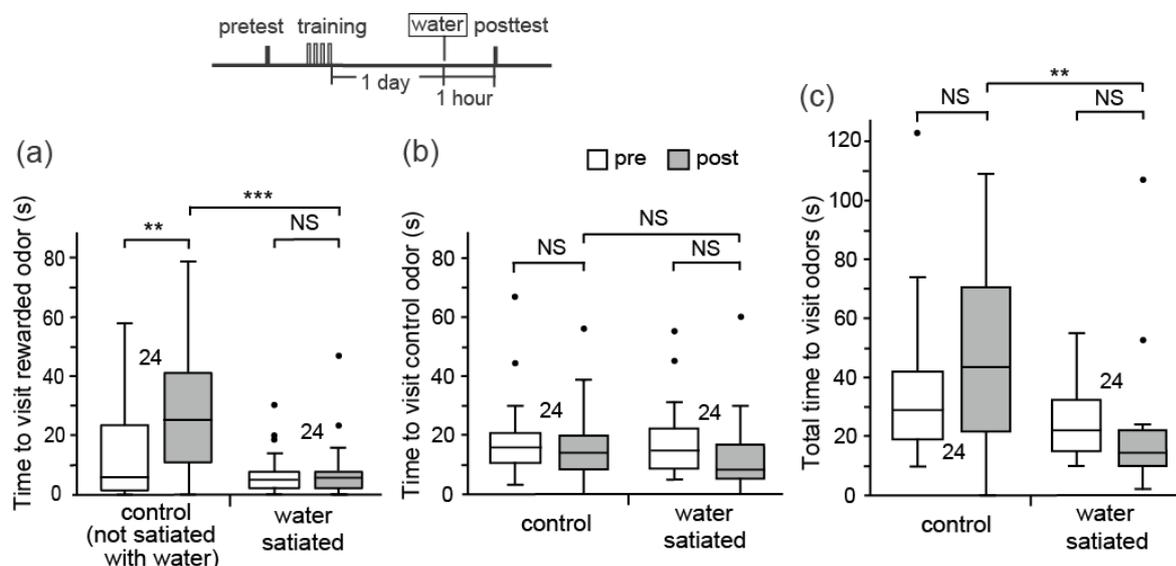


Figure S1. Effects of water satiation on responses to water-associated odors after standard training were evaluated by the time spent visiting odors. The time visiting rewarded (peppermint) odor (a) and control (vanilla) odor (b), and the total time visiting either odors (c) in the control (no water provided) group and water-satiated group shown in Fig. 1a are plotted as box plots. In this analysis, we included data from crickets that visited odor sources for less than 10 sec (two crickets for the control group and four crickets for the satiated group), which were not used for evaluation of odor preferences (see Methods). The results of statistical comparison of visiting time before and after training (WCX test) and between groups (M-W test) are shown as asterisks (* * $p < 0.01$; * * * $p < 0.001$; NS $p > 0.05$, with the p values adjusted by Holm's method). The time visiting rewarded odor was significantly longer after training than that before training in the control group but not in the devalued group (a). On the other hand, the time visiting control odor did not significantly differ before and after training in both groups (b). The total time to visiting odors did not significantly differ before and after training in both groups, but between-group comparison showed that the total visiting time after training was significantly shorter in the devalued group than that in the control group (c). The results confirm that water satiation fully suppresses responses to water-associated odor but responses to control odors are less affected.

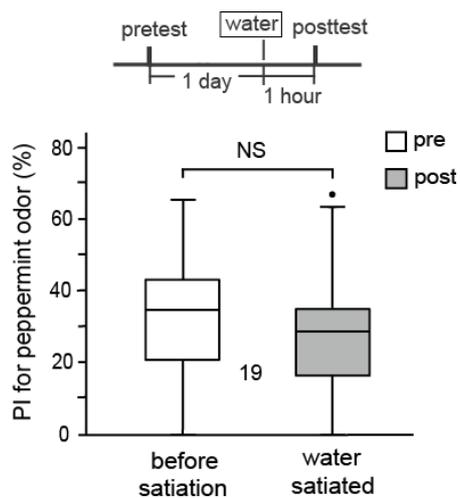


Figure S2. Water satiation does not alter odor preference in untrained crickets. Crickets in one group (n=19) were given odor preference test, and on the next day they were given water until stopped drinking, and one-hour later they were given the test again. Relative preferences between peppermint and vanilla odors, measured as the preference index (PI) for the peppermint odor, are shown as box plots. The odor preference after water satiation did not significantly differ from that before satiation (NS, $p > 0.05$, WCX test), indicating that water satiation does not alter odor preference in untrained crickets.

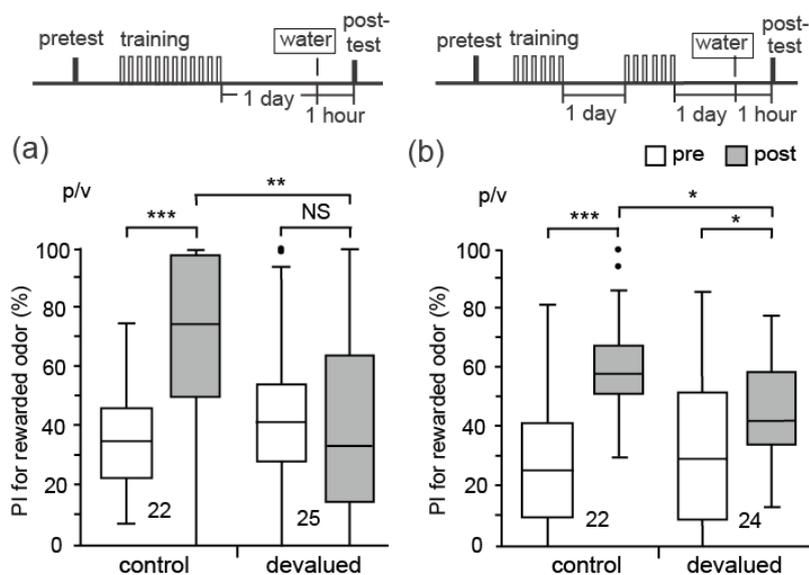


Figure S3. Evaluation of the amount of training necessary to make execution of the CR insensitive to reward devaluation. (a,b) Two groups were each subjected to a pre-test and then 12 trials x 1 day training (a), and another two groups of animals were each subjected to a pretest and then 6 trials x 2 days training (b). On the next day, crickets in one group were given water prior to the posttest and crickets in the other group were not given water. Relative preferences for the rewarded odor before (white box) and 1 day after training (gray box) are shown as box plots. The results of statistical comparison before and after conditioning (WCX test) and between groups (M-W test) are shown as asterisks (* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$; NS $p > 0.05$, with the p values adjusted by Holm's method).

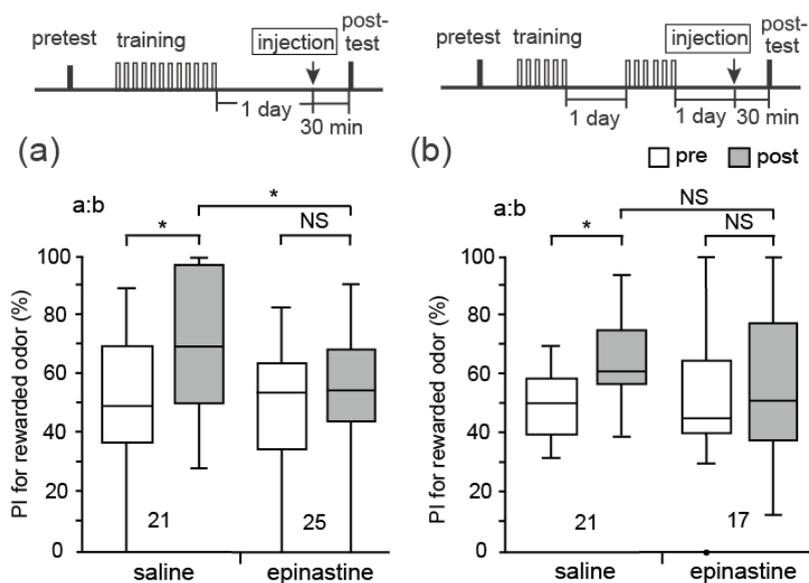


Figure S4. Evaluation of the amount of training necessary to make execution of the CR insensitive to epinastine. (*a,b*) Effects of epinastine on the CR were tested in the 12 trials x 1 day training group (*a*) and in the 6 trials x 2 days training group (*b*). Either apple odor or banana odor was used as CS and the other was used as the control odor. Relative preferences for the rewarded odor (PI) before (white box) and 1 day after training (gray box) are shown as box plots. The results of statistical comparison before and after conditioning (WCX test) and between groups (M-W test) are shown as asterisks (* $p < 0.05$; NS $p > 0.05$, with the p values adjusted by Holm's method).